Biodiversity of Magnetotactic Bacteria in the tropical marine environment of Singapore revealed by metagenomic analysis

Shi Ming Tan[1=], Muhammad Hafiz Ismail[1=], Bin Cao[1,2*]

[1] Singapore Centre for Environmental Life Sciences Engineering (SCELSE), Nanyang Technological University (NTU), 60 Nanyang Drive, SBS-01N-27, Singapore 637551

[2] School of Civil and Environmental Engineering, Nanyang Technological University (NTU), 50 Nanyang Ave, N1-01C-69, Singapore 639798

[=]Shi Ming Tan and Muhammad Hafiz Ismail contributed equally to this work.

*Corresponding author: Dr. Bin Cao

Telephone: +65-67905277

E-mail: bincao@ntu.edu.sg

Keywords: Magnetotactic Bacteria, Metagenomics, Magnetotaxis, Magnetosomes

Highlights

- This is the first report on MTB biodiversity in the tropical marine environment near the geo-equator.
- MTB abundance of ~0.2-1.69% was detected in the tropical marine environment of Singapore.
- Majority of the detected MTB represents novel taxa that cannot be classified at the species level.
- 3 MTB genomes were recovered with *Magnetovibrio blakemorei* being the closest-associated genome.
- Genera *Magnetovibrio* and *Desulfamplus*, but not *Magnetococcus*, were the dominant MTB in the tropical marine environment.

Abstract

Most studies on the diversity of magnetotactic bacteria (MTB) have been conducted on samples obtained from the Northern or the Southern hemispheres. The diversity of MTB in tropical Asia near the geo-equator, with a close-to-zero geomagnetic inclination, weak magnetic field and constantly high seawater temperature has never been explored. This study aims to decipher the diversity of MTB in the marine environment of Singapore through shotgun metagenomics. Although MTB has been acknowledged to be ubiquitous in aquatic environments, we did not observe magnetotactic behaviour in the samples. However, we detected the presence and determined the diversity of MTB through bioinformatic analyses. Metagenomic analysis suggested majority of the MTB in the seafloor sediments represents novel MTB taxa that cannot be classified at the species level. The relative abundance of MTB (~0.2-1.69%) in the samples collected from the marine environment of Singapore was found to be substantially lower than studies for other regions. In contrast to other studies, the genera *Magnetovibrio* and *Desulfamplus*, but not *Magnetococcus*, were the dominant MTB. Additionally, we recovered 3 MTB genomic bins that are unclassified at the species level, with *Magnetovibrio blakemorei* being the closest-associated genome. All the recovered genomic bins contain homologs of at least 5 of the 7 *mam* genes but lack homologs for *mamI*, a membrane protein suggested to take part in the magenetosome invagination. This study fills in the knowledge gap of MTB biodiversity in the tropical marine environment near the geo-equator. Our findings will facilitate future research efforts aiming to unravel the ecological roles of MTB in the tropical marine environments as well as to bioprospecting novel MTB that have been adapted to tropical marine environments for biotechnological applications.

1. Introduction

Magnetotactic bacteria (MTB) biomineralize iron to magnetite (Fe_3O_4) and/or greigite (Fe_3S_4) crystals in intracellular organelles called magnetosomes (Bazylinski and Frankel, 2004; Jogler and Schuler, 2009). Magnetosomes are organized into one or more chains that maximize the cellular magnetic dipole moment and subsequent responsiveness of MTB to magnetic field lines (Frankel et al., 1997; Schuler, 1999; Bazylinski and Frankel, 2004). MTB can assimilate several orders of magnitude more (about 100- to 1000- fold) ferrous and ferric ions (2-3% iron per cell by dry weight) than non-MTB cells (Heyen and Schuler, 2003), contributing substantially to the biogeochemical cycle of iron in aquatic ecosystems (Amor et al., 2020).

MTB are generally considered to be ubiquitous in sediments and chemically stratified water columns of aquatic ecosystems including freshwater, brackish, and marine environments such as lagoons, estuaries, lakes, rivers, hydrothermal vent, seamounts, and extreme environments (Bazylinski and Frankel, 2004; Lin et al., 2012b; Abreu et al., 2016; Liu et al., 2017). Vertical stratification of chemicals, *e.g.*, iron, sulfide, and oxygen, influences the distribution of MTB (Simmons et al., 2004), where MTB migrate to an optimum environmental niche using a combination of magnetotaxis, chemotaxis, and aerotaxis (Lins and Bazylinski, 2009). Large populations of MTB are often present in chemically stratified environments, especially at the oxic-anoxic interface (OAI) of sediments or water columns (Spring et al., 1993; Lin et al., 2009).

Previous studies using the 16S rRNA gene to reveal the phylogenetic diversity of MTB have shown that MTB are highly diverse, ranging from the *Alphaproteobacteria* , *Deltaproteobacteria* (Lefevre et al., 2011b) and *Gammaproteobacteria* classes (Simmons et al., 2004) of the *Proteobacteria* phylum; the *Nitrospirae* phylum (Lin et

al., 2012a; Kolinko et al., 2013); the *Candidatus Omnitrophica* phylum (Kolinko et al., 2012) and the candidate phylum *Latescibacteria* (Lin and Pan, 2015). Quite often, the *Alphaproteobacteria* is the most abundant MTB detected in aquatic ecosystems (Spring et al., 1995; Lin et al., 2014). Intriguingly, most studies on diversity of MTB have been conducted on environmental samples obtained from the Northern hemisphere such as China (Lin et al., 2012b; Lin et al., 2013), Europe (Flies et al., 2005b), and USA (Simmons et al., 2004) or the Southern hemisphere such as Rio de Janeiro, Brazil (Keim et al., 2005), Australia (Blakemore et al., 1980; Lin et al., 2018) and Antarctica (Abreu et al., 2016). In contrast, the diversity of MTB in tropical Asia near the geo-equator, e.g., the marine environment of Singapore, has not been explored. Whether and how the close-to-zero geomagnetic inclination, i.e., the angle between the magnetic field lines and the Earth surface (NOAA, 2019b), the weak magnetic field (NOAA, 2019a), and the constantly high water temperature in the region influence MTB biodiversity remains elusive.

The objective of this study was to reveal the MTB diversity in the tropical marine environment of Singapore, using marine sediments collected from the Johor Strait that separates Peninsular Malaysia and Singapore. This is the first study that deciphers the biodiversity of MTB in the tropical, near the geo-equator marine environment of Singapore.

2. Materials and Methods

2.1. Sampling and laboratory incubation of marine sediments and seawater

Sampling sites and the overall experimental design flowchart are shown in Figure 1. Samples were collected from 3 locations around the Johor Straits. Sites: 1, 2 and Sbw from the Johor Straits were specifically chosen for metagenomic sequencing

because it is less influenced by strong tidal currents and reduced tidal mixing (Tan et al., 2016) and were more likely to have chemically stratified aquatic environments that support the growth of MTB. Sediments were sampled using a Van Veen grab, where the first few centimetres of the grabbed sediments were transferred to sterile polycarbonate bottles (Treforest, USA). Seawater was collected from the respective sites using a siphon pump or a Niskin bottle mounted on a rosette system (Sea-Bird Scientific, USA). Polycarbonate bottles were filled to 1/3 of their volume with marine sediments, and the remaining 2/3 were filled with seawater. Samples were brought back to the laboratory immediately after sampling and stored at room temperature (~ 25^{o}C) and in the dark for several weeks to months, with the bottle screw cap loosely closed to allow chemical stratification and enrichment of MTB (Oestreicher et al., 2012). Each sampling site had 2 technical replicates, except for site Sbw that had three replicates: 2 technical replicates with an incubation period of 153 days (sample Sbw_153) and another replicate with an incubation period of 74 days (sample Sbw_74). Sample Sbw_74 was used as an initial proof-of-concept sequencing experiment to test for the presence of MTB.

2.2. Enrichment of magnet-responsive bacterial cells

Neodymium ring magnets (2 × 1 × 0.6 cm) were placed 1 cm above the sediment/water interface on the exterior of bottles housing the slurry samples. Magnets with opposite magnetic polarity were placed on opposite side of the bottle. For visualization of magnet-responsive bacteria, approximately 1 mL of water/sediment slurry samples adjacent to each magnet was withdrawn using Pasteur pipettes into a 1.5 mL centrifuge tube (Eppendorf, USA). Ten µL of sample was observed for magnetotaxis using the hanging drop method (Schuler, 2002) under bright field microscopy (Axio Imager Widefield Fluorescence Microscope, Zeiss, Germany). Briefly, a magnetic

stirrer was placed at the edge of the sample droplet and swimming behaviour of cells towards the magnet/droplet edge was observed with the microscope.

2.3. Genomic DNA extraction

Genomic DNA was extracted from approximately 0.5 g of the slurry samples of sediment/seawater using the FastDNA SPIN KIT for Soil (MP Biomedicals, USA), where homogenisation was performed with a vortex adapter (Qiagen, USA) at full speed for 20 minutes. Eluted DNA was further purified using genomic DNA Clean & Concentrator (Zymo Research, USA).

2.4. Metagenomic sequencing and library construction

Samples had 2 technical replicates (denoted by "_R1" and "_R2") for metagenomic sequencing, except for sample Sbw_74 which had only 1 replicate as it was an initial proof-of-concept metagenome sequencing for the presence of MTB. The Illumina HiSeq Rapid SBS Kit v2 sample (Illumina, USA) preparation protocol was used for library preparation that produced reads up to 2 x 250 bp paired-end reads. Genomic sequencing of samples was performed at the Singapore Centre for Environmental Life Sciences Engineering (SCELSE), on an Illumina HiSeq 2500 (Illumina, USA) platform. A total of 7 metagenomes were generated.

2.5. Adapter- and quality-trimming of metagenomics reads

Illumina adapters were removed, and the resulting reads were quality-trimmed at both ends to a Phred score of Q15 and a minimum length of 50 bp using the software BBDuk tools (BBMap- Bushnell B. – http://sourceforge.net/projects/bbmap) with the following settings: qtrim=rl; trimq=15; ktrim=r; k=31, mink=11; ziplevel=9; hdist=1; tpe; tbo; chastityfilter=t; minlen=50. The quality of the raw and quality -

trimmed sequences was verified using FastQC (Andrews, 2010). Trimmed reads were used for subsequent bioinformatics analyses.

2.6. Metagenomics 16S rRNA gene analysis

Sequencing reads belonging to the hypervariable regions of the SSU gene in metagenomics dataset were extracted and subsequently taxonomically classified with Metaxa2 software (Bengtsson-Palme et al., 2015). Metaxa2 software allows the classification of SSU rRNA sequencing reads down to the genus and/or species level. The SILVA SSU Ref NR 99 database Release 132 (Quast et al. (2013) was built using the Metaxa2 Database builder software (Bengtsson-Palme et al., 2018). The SILVA reference database with the following settings: --filter_uncultured T, --correct_taxonomy T, --mode divergent, --cutoffs “0,60,70,75,85,90,97”, –plus T, auto_rep T was selected as it provided the highest accuracy (proportion of correctly classified sequence segments) at the genus level. For sample comparison, the raw count for the number of matched reads for each metagenome was normalized to the library size (total number of reads) for the corresponding metagenome. An OTU is defined as a distinct taxonomic unit from the abundance matrix. The abundance matrix generated from Metaxa2 was analysed using different software for downstream processing with the Ampvis2 software (Andersen et al., 2018) and these include: (1) merging of replicate samples, (2) filtering of MTB species, plotting of (3) rarefaction curves, (4) boxplot, (5) heatmap and (6) ordination plot. Relative abundance of an OTU is calculated by dividing the number of reads for an OTU by the total number of reads in each sample.

2.7. Genome assembly of contigs

Quality-trimmed reads from the metagenomes were co-assembled into contigs using the Megahit software (Li et al., 2016) with the following settings: --presets meta-large. Only contigs with a minimum length of 200 bp were selected for further processing. 16S rRNA gene sequences were extracted from the co-assembly using the Barrnap software (https://github.com/tseemann/barrnap).

2.8. Taxonomic assignment of contigs to MTB reference genomes

Genomes of MTB (n=69) were downloaded from the NCBI and JGI databases, and a BLAST database was created using the makeblastdb tool (Altschul et al., 1990). Contigs with sequence similarity to the draft genomes were extracted using BLASTN with the following settings: e-value ≤ 1×10^{-10} and pident ≥ 90% as suggested by Arango-Argoty et al. (2016). The BLASTN tabular output file was parsed into MEGAN (Huson et al., 2016) for LCA classification of contigs using the "Long Read LCA algorithm" with the following settings: max expected e-value = 10^{-10}, min percent identity = 90%, followed by the mapping of reads with NCBI taxonomy using the 'nucl_acc2tax-Jul2019' mapping file.

2.9. Identification of magnetosome-containing contigs

A database of magnetosome proteins: *mamA*, *mamB*, *mamE*, *mamI*, *mamK*, *mamM*, *mamQ* was constructed from the UniProtKB database (UniProt, 2019). Contigs from the co-assembly were queried against the individual database of magnetosome proteins using the Centrifuge classifier (Kim et al., 2016), and the classification results were analyzed using the Pavian software (Breitwieser and Salzberg, 2016).

2.10. Functional annotation of contigs

Contigs that aligned closely to MTB genomes were extracted from MEGAN software and open reading frames (ORFs) were predicted and translated using the Prodigal software (Hyatt et al., 2010) with the following setting: -p meta. The amino acid sequences of the ORFs were annotated using web-based tool of (1) GhostKOALA (Kanehisa et al., 2016) and (2) GO-FEAT (Araujo et al., 2018).

2.11. Recovery of genomic bins

The co-assembly with a minimum length of 1000 bp were selected for genomic binning with MetaBAT2 (Kang et al., 2019) and CONCOCT binning (Alneberg et al., 2014), and the resulting MAGs were subsequently dereplicated and aggregated using DAS Tool (Sieber et al., 2018). The completeness and contamination level of the MAGs were verified with the CheckM software (Parks et al., 2015), and taxonomy of MAGs were assigned by the GTDB-Tk software (Chaumeil et al., 2019). SCGs of selected genomic bins were taxonomically verified using the “anvi-run-scg-taxonomy” of the Anvio software (Eren et al., 2015).

2.12. Phylogenetic analysis of genomic bins

The genomic bins 234_sub, bin.1085 and bin.1389 were used in a phylogenetic analysis with *Magentovibrio* spp. isolates, *Magnetovibrio blakemorei*, *Indioceanicola profundi* and *Denitrobaculum tricleocarpae* genomes obtained from NCBI. The analysis was performed with PhyloPhlAn 3.0 (Asnicar et al., 2020) and the core genes were downloaded from UniRef90 database when specified with the “-i" option.

2.13. Identifying proteins homologous to *mam* genes in genomic bins

The individual sets of mam genes were aligned using Clustal Omega on the EMBL-EBI website before being converted to HMM profiles using the “hhmake” function of

218 the hhsuite program package (Steinegger et al., 2019). The coding sequences of
219 each genomic bin and *Magnetovibrio blakemorei* were also converted to HMM
220 profiles, first by building multiple sequence alignments using “hhblits” before the
221 HMMs were computed using the “hhmake” option of “ffindex_apply”. Once the HMM
222 profiles were ready, pairwise alignment of the *mam* genes and genomic bins coding
223 sequence HMM profiles were performed using the global alignment mode of the
224 “hhblits” function.

225 **2.14. Diversity of microbial communities**

226 The abundance matrix generated from Metaxa2 was transposed and used to
227 calculate the distance matrix using the Bray-Curtis distance with the “vegdist”
228 function of the “vegan package” (Oksansen et al., 2019). Hierarchical clustering was
229 performed using the “hclust” function to produce a dendrogram. Non-Metric
230 Multidimensional Scaling Analysis (NMDS) based on the Bray-Curtis distance
231 measure (Bray and Curtis, 1957) of 7 samples and 18 MTB OTUs was performed
232 using the Ampvis2 software (Andersen et al., 2018). No initial data transformation
233 has been applied.

234 **2.15. Sequence data**

235 These sequence data have been submitted to NCBI under BioProject accession ID:
236 PRJNA563777. Metagenomes from the sampling sites Site 1, Site 2 and Sbw have
237 been deposited under BioSample accession ID:, SAMN12687858, SAMN12687859
238 and SAMN12687857 respectively. Quality-trimmed sequencing reads have been
239 submitted to Sequence Read Archive with accession ID: SRR10083956. The 16S
240 rRNA sequences obtained from the co-assembly have been deposited in NCBI’s
241 GenBank with the following accession numbers: KDOT01000001-KDOT01001363.

3. Results

3.1. Observation for magnetotactic behaviour from marine sediments

Collected samples that were stored for weeks were incubated with magnets near the water/sediment interface. No bacterial pellet or accumulation of cells was observed at the side of the magnets after 336-h incubation. Our observation contrasts the findings of other studies that have employed similar enrichment method where MTB cells can be seen accumulating adjacent to the magnet after 30 min (Flies et al., 2005b; Bazylinski et al., 2013; Xu, 2018).

Thus, the following scenarios were proposed: MTB were (1) not present, (2) in low abundance or (3) weakly-motile and weakly-magnetic (Sakaguchi et al., 1996). The hanging drop technique with bright field microscopy was used to check for the presence of MTB, with samples being obtained near the side of the magnets. No magnetotactic behaviour was observed at either pole of the magnet placed near the sample droplet. The observation was consistent for samples observed over a period of a year.

3.2. Metagenomics 16S rRNA gene analysis of MTB microbial community

Metagenomic sequencing and bioinformatics analyses were performed on the DNA directly extracted from the sediment/water slurries collected from the Johor Straits to reveal the biodiversity of MTB in the marine environment of Singapore (Figure 1).

(A)

(B)

Sampling site, date, and sediment color	Geographic coordinates	Sample name	Incubation period*
Sbw (28/6/2018; black)	1°28'16.3"N 103°48'16.3"E	Sbw_153_R1^	153 days
		Sbw_153_R2^	
		Sbw_153#	
		Sbw_74	74 days
Site 1 (13/11/2018; brown)	1°26'50.8"N 103°45'48.8"E	1B_R1^	58 days
		1B_R2^	
		Site1#	
Site 2 (13/11/2018; grey)	1°27'26.3"N 103°43'14.1"E	2B_R1^	21 days
		2B_R2^	
		Site2#	

* Incubation period of samples prior to DNA extraction
^ Samples with two technical replicates for metagenomic sequencing
# Sample name of merged technical replicates

(C)

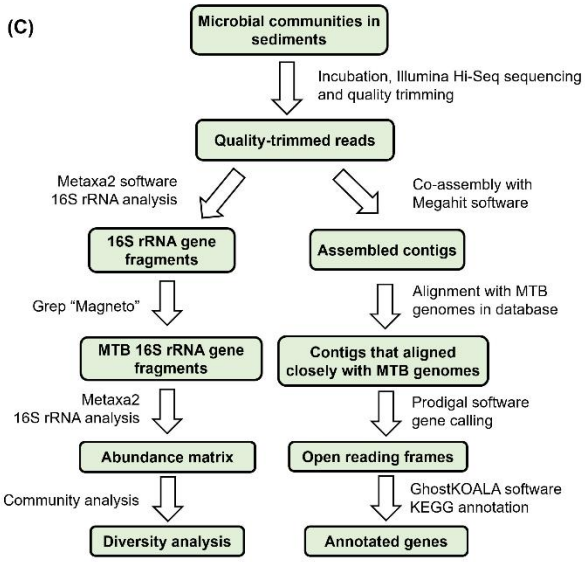

Figure 1. (A) Sampling location of sediments and seawater around the Straits of Johor. (B) Description of the sampling sites, incubation duration of samples and observation of sediment colour. (C) Flowchart illustrating the experimental design.

A total of 1,287,083,716 paired-end reads, with an average read size of 245 bp were generated from the 7 metagenomes after quality-trimming (Supplementary Table S1). To identify potential MTB from the metagenomic 16S rRNA gene analysis, the term "Magneto" was used to search the taxonomic affiliations of the extracted 16S rRNA gene sequences. A total of 22 MTB OTUs (including unclassified MTB species) were identified, and four (*unclassified Telmatospirillum*, *Varunaivibrio sulfuroxidans*, *Candidatus Nitrosoarchaeum limnia SFB1* and *alpha proteobacterium NH6-6*) were removed because a conducted literature review showed the absence of magnetotactic behaviour in these species (Table 1).

277 **Table 1**. Relative abundance of raw read counts of potential MTB as analysed by 16S rRNA gene analysis.
278

Classification						Samples (Relative abundance, %)						
Phylum	**Class**	**Order**	**Family**	**Genus**	**Species**	**1B_R1**	**1B_R2**	**2B_R1**	**2B_R2**	**Sbw_153_ R1**	**Sbw_153_ R2**	**Sbw_74**
Nitrospirae	Thermodesulfovibrionia	Magnetobacteriales	Magnetobacteriaceae	Candidatus Magnetobacterium	Unclassified	0.0027	0	0	0.0059	0.0058	0.0103	0.0017
				Candidatus Magnetoovum	Unclassified	0.0054	0.0027	0.0099	0.0147	0.0088	0.0129	0.0069
				Unclassified	Unclassified	0.0054	0.0106	0.0331	0.0383	0.0263	0.0283	0.0414
Omnitrophicaeota	Candidatus Omnitrophus magneticus	Unclassified	Unclassified	Unclassified	Unclassified	0.0054	0	0	0	0	0	0
Proteobacteria	Alphaproteobacteria	Rhodospirillales	Magnetospiraceae	Magnetospira	Unclassified	0.0054	0	0	0	0.0117	0.0103	0
				Magnetovibrio	Magnetovibrio blakemorei	0.0162	0.0186	0	0	0.4993	0.4712	0
					Unclassified	0.0325	0.0425	0	0	0.6395	0.5973	0
					Alphaproteobacterium PR-2	0.0568	0.1064	0	0	0.4526	0.3450	0
				Unclassified	Unclassified	0.0108	0.0080	0.0033	0	0.1168	0.1313	0
			Magnetospirillaceae	Magnetospirillum	Unclassified	0	0	0	0	0.0029	0.0051	0.0052
				Unclassified	Unclassified	0	0	0.0033	0.0029	0.0058	0.0026	0.0035
	Deltaproteobacteria	Desulfobacterales	Desulfobacteraceae	Candidatus Magnetananas	Unclassified	0	0	0.0033	0.0029	0	0	0.0035
				Candidatus Magnetoglobus	Candidatus Magnetoglobus multicellularis str Araruama	0	0	0.0033	0	0	0	0
					Unclassified	0	0	0	0.0029	0	0	0
				Candidatus Magnetomorum	Candidatus Magnetomorum litorale	0.0027	0	0.0331	0.0530	0.0146	0.0257	0.0035
					Unclassified	0.0027	0.0027	0.0033	0.0118	0.0029	0	0.0017
	Magnetococcia	Magnetococcales	Magnetococcaceae	Magnetococcus	Unclassified	0	0.0027	0.0033	0.0059	0	0.0026	0
					Magnetic coccus MP17	0	0	0	0	0	0.0026	0

279

MTB sequences were detected in all sampling sites, with sample Sbw_153 having the highest relative abundance of MTB (n = 14, relative abundance = 1.69 ± 0.19%) and Sample Sbw_74 having the lowest relative abundance of MTB (n = 8, relative abundance = 0.067 ± 0.010%). Hierarchical clustering (Figure 2A) and NMDS analysis (Figure 2B) grouped technical replicates into tight clusters. Since the microbial community structure of replicates was similar, read counts in replicate samples were aggregated and mean abundance for each taxonomic group was calculated. The MTB communities sampled from different locations were placed further apart, except for sample Sbw_74 that was clustered closer towards samples from Site 2.

Rarefaction curves where the number of reads were plotted against the number of observed MTB OTUs for each sample are shown in Figure 2C. Sbw_153 was the only sample nearly reaching asymptote for the rarefaction curve, indicating that it had the highest alpha diversity and had captured majority of the MTB diversity (n = 14). Coincidently, sample Sbw_153 had the longest incubation period for MTB enrichment (Figure 1B). This observation is in line with studies (Flies et al., 2005b; Lefevre et al., 2011a) which showed that prolonged incubation of microcosms over several months led to the increase in abundance of certain MTB cells. The rarefaction curves showed that majority of the MTB species in the samples have not been observed. Therefore, a greater sequencing effort is required as MTB constitute a small fraction of the sampled microbial communities. However, the high complexity of microbial communities present in sediments/soil makes sequencing-to-saturation and genome recovery of rare taxa a challenge (Gans et al., 2005; Jansson et al., 2012).

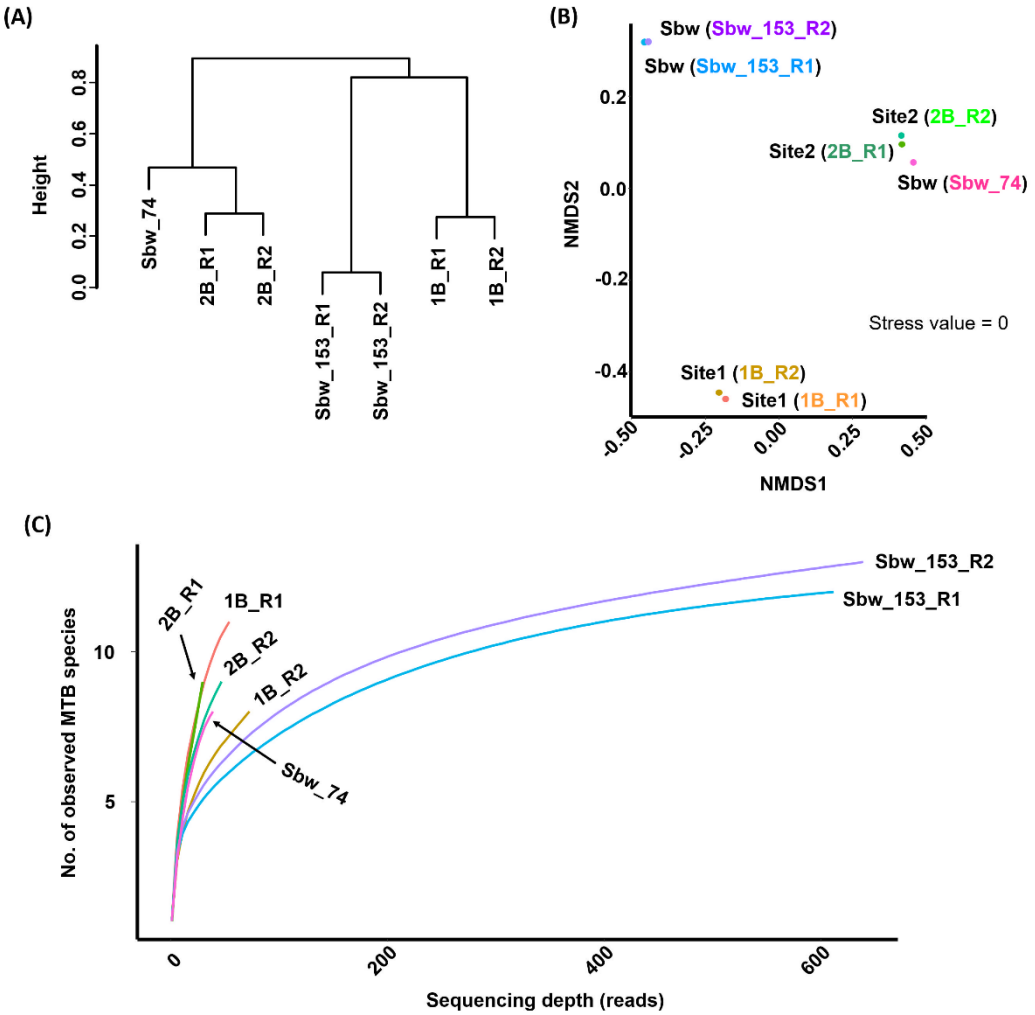

Figure 2. Comparison of MTB communities obtained from different sampling locations around the Johor Straits, Singapore using (A) average linkage hierarchical clustering and (B) non-metric multidimensional scaling (NMDS) ordination plot based on Bray-Curtis dissimilarity. Samples in the NMDS plot are labelled by their sampling sites. (C) Rarefaction curves for MTB communities.

MTB belonging to the phylum *Proteobacteria* and *Nitrospirae* were detected in all the sites, but only phylum *Omnitrophicaeota* was present in Site 1 (Figure 3A). *Candidatus Omnitrophus magneticus*, an MTB representative of the candidate phylum *Omnitrophica* that previously has not been identified in the marine environment (Kolinko et al., 2016) was detected in Site 1 (relative abundance = 0.0026%). MTB species belonging to the phylum *Nitrospirae* were detected in all the sampling sites and could be classified to the genus *Candidatus Magnetoovum* and

320 *Candidatus Magnetobacterium*. Our findings corroborate a previous study (Xu, 2018)
321 where MTB from the phylum *Nitrospirae* were detected for the first time in the marine
322 environment. Overall, *Alphaproteobacteria* is the most abundant class of MTB (n = 7
323 MTB species, mean relative abundance = 0.063 ± 0.16%), which is in agreement
324 with most biogeographic studies of MTB performed in various aquatic systems
325 (Spring et al., 1995; Lin et al., 2012b). In previous studies, the most abundant genera
326 of MTB present in environmental ecosystems was reported as the magnetotactic
327 cocci belonging to the order *Magnetococcus* within the class *Alphaproteobacteria*
328 (Flies et al., 2005b; Lefevre and Bazylinski, 2013; Lin et al., 2014; Lin et al., 2017a).
329 In contrast, only 2 out of the 18 OTUs (mean relative abundance = 0.018 ± 0.0022%)
330 of *Magnetococcus*, *Magnetic coccus* MP17 and unclassified *Magnetococcus*, were
331 identified in the present study. As 16S rRNA classification was performed using the
332 SILVA taxonomy, *Magnetococcus* was assigned to the class *Magnetococcia* (SILVA
333 taxonomy) instead of *Alphaproteobacteria* (NCBI taxonomy). In our study, the most
334 abundant MTB species is the unclassified *Magnetovibrio* (0.19 ± 0.087%) while the
335 most abundant classified MTB species is *Magnetovibrio blakemorei* (0.14% ± 0.055)
336 even though it was not detected in Site 2 (Figure 3B).

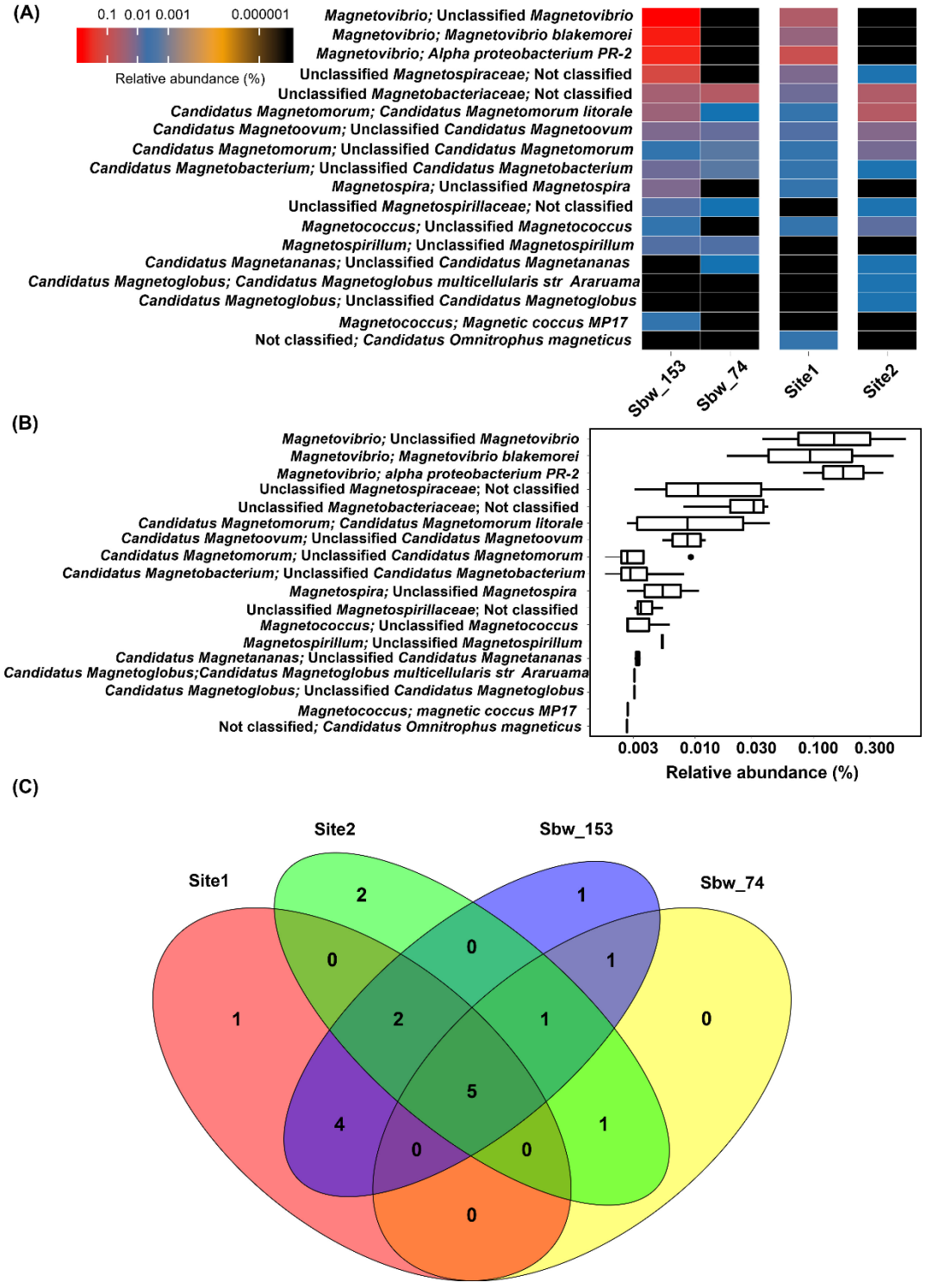

Figure 3. Metagenomic 16S rRNA analysis of MTB communities. (A) Heatmap depicting the relative abundance of MTB species at individual site. (B) Boxplot depicting the cumulative relative abundance of MTB species across all sampling sites. (C) Venn diagram showing the number of shared MTB species among all samples.

A Venn diagram illustrating the core MTB OTUs that were shared among the samples is shown in Figure 3C, in which the relative abundance was not taken into consideration (Figure 3C). Only 5 OTUs were identified to be common among the samples, and 4 of them represent unclassified MTB species. *Candidatus Magnetomorum litorale* is the most abundant classified MTB species detected in all the sampling sites. It has been previously identified as a greigite-synthesizing multicellular magnetotactic prokaryote (MMP) from the coastal North Sea tidal sand flats (Wenter et al., 2009). Besides *Candidatus Magnetomorum litorale*, *Candidatus Magnetoglobus* (unclassified at the species level) is the only MTB OTU exclusively detected in Site 2. Species associated with the *Candidatus Magnetoglobus*, such as *Candidatus Magnetoglobus multicellularis* (Abreu et al., 2014) and *Candidatus Magnetomorum* sp. 1–7 (Zhang et al., 2014) have been identified as gregite-producing and gregite/magnetite MMPs, respectively. *Candidatus Magnetomorum*, *Candidatus Magnetoovum*, *Candidatus Magnetobacterium* and unclassified *Magnetobacteriaceae* were found to be common MTB OTUs at the genus level detected in all our sampling sites. This finding differs from other studies that often report the genera *Magnetospira*, *Magnetococcus* or *Magnetovibrio* (Dong et al., 2016) to be the core MTB OTUs present in sea surface sediments.

In summary, we were able to detect MTB species belonging to the *Alphaproteobacteria*, *Deltaproteobacteria*, *Nitrospirae* and *Omnitrophicaeota*, and the same taxonomic lineage could be found in other large-scale study of MTB from varied aquatic environments from the Northern and Southern hemisphere (Lin et al., 2018). The diversity of MTB obtained in this study is greater than other marine biogeographic studies that have reported the diversity of MTB. For instance, Xu (2018) reported MTB OTUs affiliated to *Alphaproteobacteria*, *Deltaproteobacteria*

and *Nitrospirae* in sediments of the Yellow Sea, China, and Dong et al. (2016) reported MTB OTUs affiliated to only the *Alphaproteobacteria* and *Deltaproteobacteria*.

The percentage identity of each MTB rRNA sequence to its best BLAST match in the SILVA database was plotted against its taxonomic classification on the class level and coloured based on its taxonomic classification at the species level (Figure 4). Majority of the detected MTB species were either: (1) unclassified MTB species (9/18 OTUs, mean = 0.021 ± 0.052%) or (2) MTB that could not be classified at the species level (3/18 OTUs, mean = 0.022 ± 0.017%). The median percentage identity of all the MTB OTUs was 96.20% and the lowest percentage identity was 80.74%.

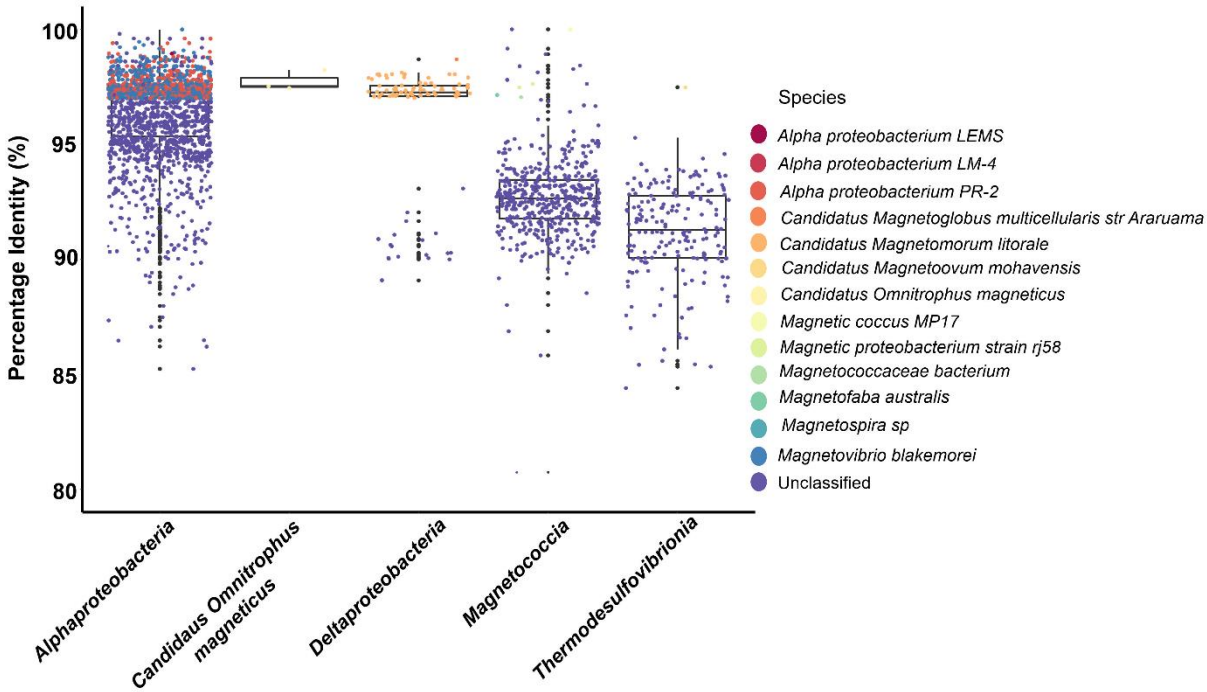

Figure 4. Percentage identity of MTB 16S rRNA sequence to its best BLAST match in the SILVA database plotted against its taxonomic classification on the class level, and coloured based on its taxonomic classification at the species level.

Intriguingly, the biodiversity of MTB for Sbw_153 differed from Sbw_74, thus suggesting that incubation of the sediment/water slurries to (re)establish a stratified environment influences MTB growth and/or detection. Seven MTB OTUs (unclassified *Magnetospiraceae*, unclassified Magnetospira, unclassified *Magnetovibrio*, unclassified *Magnetococcus*, *Magnetovibrio blakemorei*, Alpha-proteobacterium PR-2, and magnetic coccus MP17) were observed only in sample Sbw_153 after the prolonged incubation and were not present in sample Sbw_74 (Supplementary Figure S1). It is plausible that MTB communities are susceptible to the changes in the vertical physicochemical gradients from temporal variation in the microcosms. For instance, magnetotactic coccoid bacteria are often present in aged microcosms, and MMPs are often present in fresher microcosms (Flies et al., 2005a; Postec et al., 2012). While prolonged incubation for Sbw_153 led to certain MTB species being detected as shown with the 7 MTB OTUs, one single MTB OTU (*unclassified Candidatus Magnetananas*) was detected in Sbw_74 but not in Sbw_153. This observation is in line with the observation of Martins et al. (2012) that described the decline in abundance of *Candidatus Magnetoglobus multicellularis*, which belongs to the same family, Desulfobacteraceae, as *unclassified Candidatus Magnetananas*, over time in microcosms. The MTB community dynamics during the incubation of the microcosms and its environmental controls would be an interesting topic for further investigation in the future.

3.3. Taxonomic annotation of contigs

Genomic co-assembly was performed on the metagenomes (n = 7) and statistics of the assembly is presented in Supplementary Table S2. As compared to other studies that have sequenced complex microbial communities from soil samples, our study employed a higher sequencing depth to target the low abundant MTB. A total of

1,887 16S rRNA gene sequences were retrieved from the co-assembly of contigs. Among them, 214 chimera sequences, and 310 sequences deemed to have low or no similarity to 16S rRNA gene or were misassembled as detected by NCBI's Ribosensor tool were removed. Of the remaining 1,363 sequences, 321 sequences were unclassified even at the Kingdom level and 1,187 sequences were unclassified at the genus level. Two sequences with the length of 657bp and 429 bp (accession numbers: KDOT01000077 and KDOT01000735) could be taxonomically classified to the genus *Magnetovibro* with a sequence identity of 95.45% and 91.22%, respectively.

To recover contigs whose sequences are highly similar to MTB, contigs were aligned using BLASTN against a reference database containing 69 MTB draft genomes (Supplementary Table S3). The contigs were assigned to 45 distinct MTB genomes based on the following BLASTN options: evalue < 1e-10; pident > 90%; length > 100 bp (Arango-Argoty et al., 2016). From the BLASTN results, contigs were grouped by their distinct taxonomy and contigs with the greatest alignment length and highest bitscore were filtered out. Subsequently, these contigs were ranked in accordance to their alignment length and taxonomy at the species level, and the bitscore of the aligned contigs is also shown (Figure 5A).

The longest alignment length of the contig from the co-assembly to the database of MTB genomes was 8,599 bp, with a bitscore of 11,256 bp and a sequence similarity of 90.36% to the metagenome-assembled bin *Magnetovibrio sp*. Isolate SP125 (Tully et al., 2018). This finding is consistent with the 16S rRNA gene analysis that showed the genus *Magnetovibrio* ranked as the most abundant MTB species from the sampling sites (Figure 3B). A contig alignment of 1,953 bp, with a bitscore of 2,558 was found to match to *Magnetovibrio blakemorei* with a sequence similarity of

90.37% and this MTB species was identified as the most abundant classified MTB species based on 16S rRNA gene analysis (Table 1). Interestingly, the second longest alignment of 2,134 bp, with a bitscore of 2,748 matched to *Desulfamplus magnetovallimortis* with a sequence similarity of 90.16%, although *Desulfamplus magnetovallimortis* was not detected in the metagenomics 16S rRNA gene analysis (Figure 3A). This is likely because the 16S rRNA gene sequence of *Desulfamplus magnetovallimortis* is not included in the SILVA SSU Ref NR 99 database Release 132, highlighting the usefulness of shotgun metagenomics for detection of other gene fragments besides the 16S rRNA marker gene.

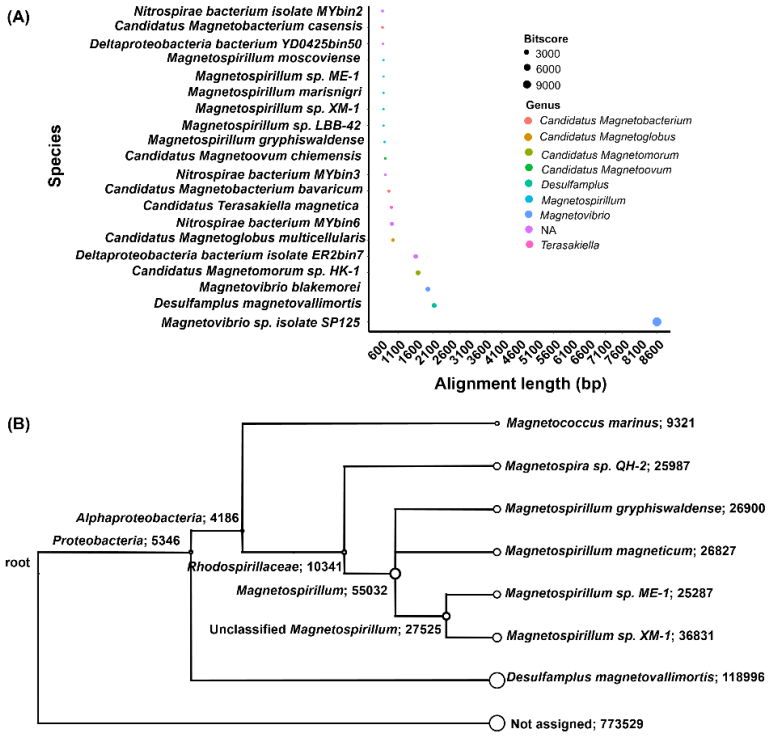

Figure 5. Taxonomic classification of contigs based on BLASTN alignment against a database of 69 MTB draft genomes using BLASTN. (A) Contigs were grouped by their distinct taxonomy, and contigs with the highest alignment length and bitscore were filtered out. The alignment length of the filtered contigs were plotted against their taxonomic classification at the species level. The colours of the plotted points represent taxonomic classification at the genus level, and the size of the plotted points is directly proportional to the bitscore of the contig. (B) Taxonomic classification of aligned contigs from the BLASTN output using the lowest common ancestor (LCA) algorithm of MEGAN software.

BLASTN results were parsed into MEGAN software (Huson et al., 2016), and the lowest common ancestor (LCA) algorithm showed *Magnetospirillum* as the most abundant MTB genus. All the contigs were taxonomically classified to the phylum *Proteobacteria* (Figure 5B), supporting other studies that show magnetotaxis is widespread in the phylum *Proteobacteria* and the majority of MTB in aquatic ecosystems belongs to this phylum (Lin et al., 2018). *Desulfamplus magnetovallimortis* is the MTB species with the most number of hits (44.05%) (Supplementary Figure S2). *Desulfamplus magnetovallimortis* has been described as a strict anaerobe with a genome that consists of two sets of magnetosome gene clusters (MGCs) capable of synthesizing greigite and/or magnetite nanocrystals (Lefevre et al., 2011b; Descamps et al., 2017). Contigs taxonomically classified to MTB at the species level were extracted (n = 1103) for subsequent functional annotation.

3.4. Identification of contigs containing *mam* genes

Taxonomic classification can also be performed with the amino acid sequences of homologous MTB-specific proteins, such as the *mam* (magnetosome membrane) genes, which are organized in close proximity as MGCs (Lefevre and Bazylinski, 2013) or magnetosome gene island (MAI) (Ullrich et al., 2005) that contain features (e.g. transposase genes) that flank the *mam* genes (Schuler, 2008). Previously, homologs of core magnetosome genes from MGCs have been used to retrieve contigs that belong to magnetosome genes from public genomic databases (Lin et al., 2017b). Contigs containing the magnetosome genes can be phylogenetically assigned, and the environments from where these contigs originated from can be included in the biogeographical studies of MTB (Lin et al., 2017a). Based on phylogenetic analysis of concatenated magnetosome protein amino acid sequences,

481 novel MTB that that shared a close phylogenetic relationship with
482 *Gammaproteobacteria* MTB strain SS-5 were discovered in marine oxygen minimum
483 zones (Lefevre et al., 2012). Similarly, MGC-containing contigs were discovered from
484 novel candidate phylum *Latescibacteria* (Lin and Pan, 2015) and *OP3* (Kolinko et al.,
485 2012), thus further expanding our knowledge on the diversity of MTB.

486 The same approach was applied in this study where a database of a core set of
487 orthologous genes shared by *Alpha*- and *Deltaproteobacteria* MTB that biomineralize
488 gregite and magnetite, i.e., *mamA*, *mamB*, *mamE*, *mamI*, *mamK*, *mamM*, *mamQ*
489 *(Lefevre et al., 2013)*, was constructed from the UniProt database (UniProt, 2019).
490 Contigs from the co-assembly were screened against this database using a novel
491 indexing scheme based on the Burrows-Wheeler transform (BWT) and the
492 Ferragina-Manzini (FM) index with the Centrifuge microbial classification software
493 (Kim et al., 2016). The *mam* genes were identified among the contigs and their
494 corresponding taxonomic classifications are presented in Supplementary Table S4.
495 The *mamE* gene had the highest number of hits (n = 596) and *mamQ* had the lowest
496 number of hits (n = 21). *Gammaproteobacterium* SS-5 and *Deltaproteobacterium*
497 ML-1 are the only MTB species that were found to contain all of the *mam* genes
498 (Supplementary Figure S3). This finding contrasts the 16S rRNA metagenomic gene
499 and MTB genome analysis, where *Gammaproteobacteria* was not detected because
500 a representative genome is lacking in the database.

501 To further analyze the *mam* genes classification, BLASTN was carried out and *mam*
502 genes could be taxonomically classified to *Gammaproteobacterium* SS-5. However,
503 classification of *mam* genes in the metagenomics dataset should be interpreted with
504 caution because most of the BLAST alignments length was substantially shorter than
505 the length of the *mam* genes and homology search was only based on partial

506 sequence of the genes (Supplementary Figure S4). The short alignment length of
507 contigs to *mam* genes suggest that it was not possible to obtain complete MGCs
508 from the metagenomes to study gene content and organization of MGCs.
509 Furthermore, all the reads were classified as “not assigned” using the default LCA
510 parameters when BLAST results were parsed into MEGAN software. This shows that
511 the sequence alignment by BLAST did not meet the MEGAN software required
512 threshold for a hit to be considered as “classified”.

513 **3.5. Functional annotation of contigs**

514 Analysis of MTB genomes could yield insights into molecular adaption of the MTB
515 community to ecological niches. A total of 1,708 ORFs were predicted from the
516 extracted contigs that aligned closely to MTB genomes and were functionally
517 annotated. Two giant genes (>5000 amino acids) were present and one could be
518 annotated to *recQ* gene (KO3654) that encodes for ATP-dependent DNA helicase
519 RecQ. A total of 41 genes had a length of over 2000 amino acids (Supplementary
520 Table S5), and some of the notable annotated genes were *cheA* gene (K03407) and
521 methyl-accepting chemotaxis protein (K03406).

522 Nitrogenase genes *nifB* gene (K02585) and *nifU/iscU* gene (K04488) responsible for
523 nitrite reduction were detected in the metagenomes (Supplementary Table S6). The
524 potential involvement of the MTB in the Johor Straits in sulfate reduction is plausible
525 as the *dsrB* gene (K11181) encoding for sulfite reductase beta subunit was
526 identified. Other detected genes involved in sulphur-metabolism included the *aprB*
527 gene (K00395) for adenylylsulfate reducate subunit B enzyme, *qmoA* gene (K16885)
528 for quinone-modifying oxidoreductase subunit QmoA and *cysK* gene (K01738) for
529 cysteine synthase. Since MTB accumulate more iron than non-MTB, most MTB

genomes contain genes for iron acquisition system such as the Feo-dependent ferrous transport system located within the MGI, Tpd-like high-affinity ferrous iron transporter genes, Fe(III)-siderophore complexes and TonB-dependent receptors. However, we were not able to locate genes involved in iron uptake except for *afuB*/*fbpB* gene (K02011) that encodes for iron (III) transport system permease protein.

Interestingly, a large number of transmembrane chemoreceptors genes encoding for methyl accepting chemotaxis proteins (MCPs) that are responsible for chemotaxis and magnetotaxis were detected, including *cheA* (K03407), *cheB* (K03412), *cheR* (K00575), *cheW* (K03408), *cheY* (K03413), *motA* (K02556), *motB* (K02557) and a MCP protein where most of the contigs were functionally annotated to (K03406) (Supplementary Table S6). MCPs sense environmental signals and modulate the flagellar motor switch protein and subsequent direction of the flagellar motor through a CheA-CheY signal transduction system (Falke and Hazelbauer, 2001; Hazelbauer et al., 2008). MTB genomes often contain a high number of genes for MCPs. For example, there are more than 60 MCPs encoded in the genomes of *Magnetospirillum magnetotacticum* strain MS-1 (Alexandre et al., 2004) and *Magnetospirillum magneticum strain AMB-1* (Matsunaga et al., 2005). Among the annotated contigs, 3 genes belonging to the gene ontology cyclic-di-GMP binding (GO:0035438) which contain the PilZ domain-containing protein were present (Supplementary Table S7), and these genes have been reported to interact selectively and non-covalently with cyclic-di-GMP (Ryjenkov et al., 2006). The presence of PilZ domain-containing proteins, which has shown to regulate bacterial biofilm formation, flagellum-driven motility and virulence (Pratt et al., 2007; Baker et al., 2016) suggests a possible biofilm mode of life for MTB. The presence of one

555 gene (GO:0004888) encoding for protein BdlA, which contains domains that respond
556 to environmental signals and catalyzes biofilm dispersion in *Pseudomonas*
557 *aeruginosa* (Morgan et al 2006) further supports the hypothesis for a biofilm mode of
558 life for MTB, which warrants further investigation.

559 **3.6. Recovery of MTB genomes**

560 Genome binning was performed on the co-assembly, and the genomic bins were
561 taxonomically classified using the GTDB-Tk software. Out of 277 genomic bins
562 produced from the DAS Tool software, 5 genomic bins could be taxonomically
563 affiliated with MTB: 1 to the family *Magnetovibrionaceae*, 1 to the family
564 *Magnetospiraceae* and 3 to the genus *Magnetovibrio*. None of the genomic bins
565 could be classified at the species level. To further verify the taxonomy of the
566 genomic bins, single-copy core genes (SCGs) from the respective genomic bins
567 were analysed and compared to the SCGs of genomes obtained from the Genome
568 Taxonomy Database. Only 3 out of the 5 genomic bins had SCGs associated with
569 *Magnetovibrio blakemorei*, and the SCGs of the genomic bins have a percentage
570 identity of less than 97% to *Magnetovibrio blakemorei* (Table 2).

571 Phylogenetic analysis of these 3 genomic bins, *Magnetovibrio blakemorei*, other
572 *Magnetovibrio* spp. isolates, and 2 unrelated *Rhodospirillaceae*, using core genes of
573 either *Magnetovibrio blakemorei* or *Magnetospirillum magneticum* obtained from
574 UniRef90 showed the 3 bins and *Magnetovibrio blakemorei* clustering in the same
575 clade (Figure 6A and B). When the phylogenetic tree was generated using the *mam*
576 genes as core genes, complete trees could not be obtained (Supplementary Figure
577 S5).

As an alternative, homologs of *mam* genes were searched for using pairwise alignments of hidden Markov models (HMMs) with the protein sequences of the 3 genomic bins. All bins have homologs of at least 5 of the 7 *mam* genes, with all bins lacking homologs for *mamI*, a membrane protein suggested to take part in the magenetosome invagination (Barber-Zucker et al., 2016) (Figure 6C and Supplementary File S5).

584 **Table 2.** Taxonomic classification of MTB genomic bins based on GTDB-Tk and SCGs.

BinID	GTDB-Tk Classification	Completion (%)	Contamination (%)	Genome size (Mbp)	SCG Classification	Identity to SCG (%)
234_sub	p_Proteobacteria;c_Alphaproteobacteria;o_Rhodospirillales_A;f_Magnetovibrionaceae;g_Magnetovibrio;s_unclassified	68.42	5.33	2.64	s_Magnetovibrio blakemorei	94.2
bin.1389	p_Proteobacteria;c_Alphaproteobacteria;o_Rhodospirillales_A;f_Magnetovibrionaceae;g_Magnetovibrio;s_unclassified	95.85	4.19	3.39	s_Magnetovibrio blakemorei	96.6
bin.1085	p_Proteobacteria;c_Alphaproteobacteria;o_Rhodospirillales_A;f_Magnetovibrionaceae;g_Magnetovibrio;s_unclassified	97.51	6.07	3.26	s_Magnetovibrio blakemorei	93.7

585

586

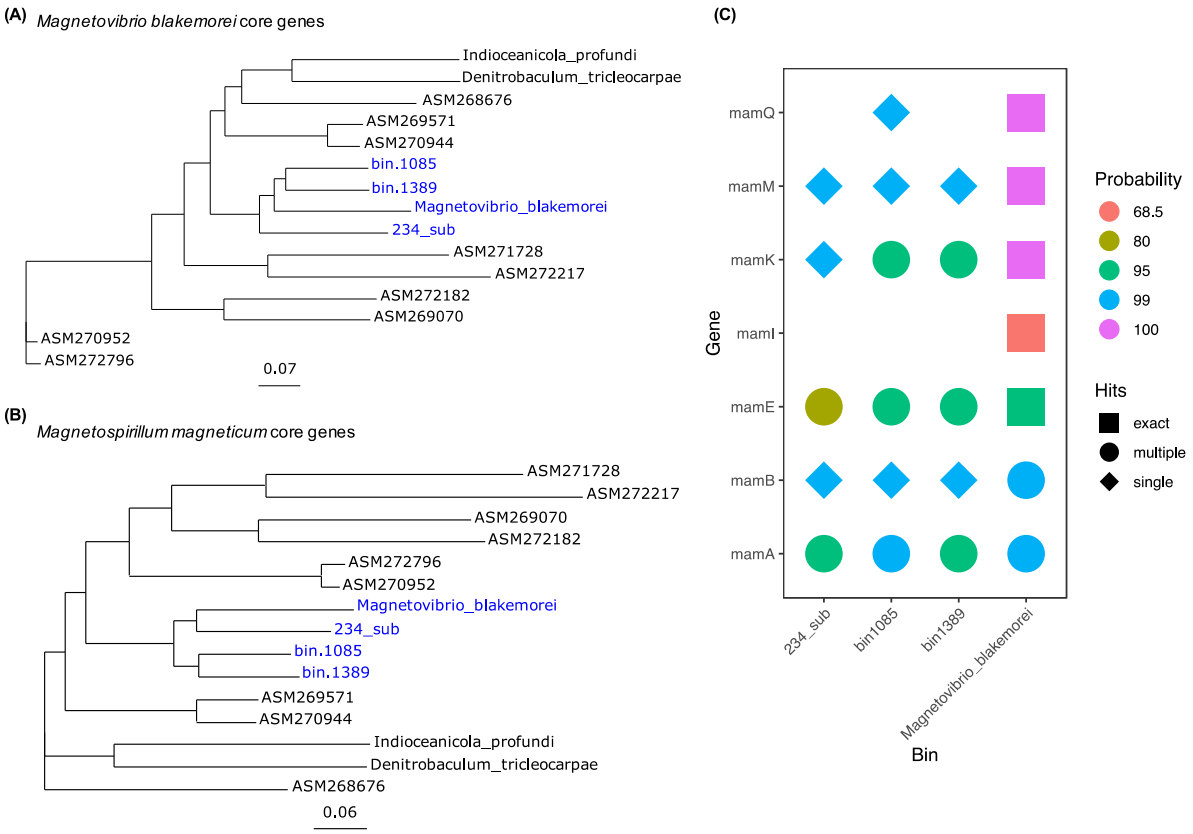

Figure 6. Phylogenetic analysis using core genes of (A) *Magnetovibrio blakemorei* and (B) *Magnetospirillum magneticum* of MTB bins. Genomes of *Magnetovibrio* spp. isolates and two other *Rhodospirillaceae*: *Indioceanicola profundi* and *Denitrobaculum tricleocarpae* were included as comparisons. (C) Homologs to seven established *mam* genes were scanned in the MTB bins using pairwise alignment of Hidden Markov Models (HMMs). The output of the analysis either showed exact matches or having 1 or more possible homologs to the *mam* genes in the genome. The shapes were coloured by the probability of the query being homologous to the template HMMs.

4. Discussion

While large-scale MTB biogeography studies have been performed in the Northern and Southern Hemisphere, studies of MTB in tropical marine environments near the geo-equatorial region are lacking. Singapore is located in the equatorial region and is composed of two main coastal bodies: the Singapore Straits in the South and Johor Straits in the North that are exposed to two monsoon seasons. Previous metagenomic sequencing performed on samples from the Johor Strait have shown that microbial communities in the Johor Strait are largely influenced by changes in salinity rather than monsoon seasonal patterns (Chénard et al., 2019). While MTB has been accepted to be ubiquitous in the sediments of freshwater and marine habitats (Flies et al., 2005a; Faivre and Schuler, 2008), we did not observe magnetotactic behaviour observed through microscopy even after the prolonged incubation period with magnets. As such, we could not address the proportion of North-seeking and South-seeking MTB in the geo-equatorial region (Frankel et al., 1981).

To determine the diversity of MTB in our samples, direct metagenomic sequencing of the marine sediments under prolonged storage was performed to overcome any bias against MTB that are tightly associated with sediment microbial communities, or MTB that exhibit low motility and weak magnetotactic activity under the enrichment conditions (Sakaguchi et al., 1996). A prolonged incubation period for the sediment samples for sample Sbw from 74 days to 153 days increased the proportion of MTB. Therefore, it is advantageous to increase the incubation period as outlined in other MTB studies (Lefevre and Bazylinski, 2013). The relative abundance of MTB in the samples collected from the Johor Straits (~0.2-1.69%) was found to be substantially

lower than previous studies for other regions. For example, in some ecosystems, MTB were the dominant proportion of the microbial community, making up to 10-50 % of the total microbial biomass (Spring et al., 1993; Simmons et al., 2007; Lin et al., 2012b). The low abundance of MTB in our sampling sites could be explained by the small inclination angle of the Earth's magnetic field at our sampling site where magnetotaxis imparts no advantage in guiding MTB along vertical axis to favourable conditions (Bennet et al., 2014).

Metagenomic analysis based on sequence similarity of sequencing reads with 16S rRNA gene sequences of established MTB species in the SILVA database had revealed a diverse MTB community in the Johor Straits. In a similar study, Dong et al. (2016) had performed amplicon sequencing of deep sea sediments obtained from the eastern Pacific Manganese Nodule Province and showcased the diversity of the MTB communities based on sequence homology to known genera of MTB despite not observing magnetotactic behaviour. The 16S rRNA gene has been frequently used as a marker gene for the classification of MTB into the different phyla (Lefevre and Wu, 2013).

A limitation of analysing the 16S rRNA gene is that we cannot predict if the putative MTB possess the ability to synthesize magnetosomes on the basis of their 16S rRNA gene sequence similarity and fit the description of MTB possessing magnetotactic traits. Another limitation is the classification of MTB and non-MTB to the same taxonomic group at the genus level based on the similarity of the 16S rRNA gene. For instance, *Desulfovibrio magneticus* strain RS-1 and *Desulfovibrio burkinensis* can be classified to the *Desulfovibrio* genus, but only *Desulfovibrio magneticus* strain RS-1 has the magnetotactic phenotype. This requires the classification of MTB down to the species level, but a large fraction of the putative MTB could not be classified at

the species level with sufficient reliability (Bengtsson-Palme et al., 2015) and were therefore labelled as “unclassified” or were missing taxonomic classification at the species level. This could be due to the input sequences belonging to new taxa that are not present in the database, or the input sequences originating from a region of the 16S rRNA gene that does not have sufficient resolution to discriminate between closely related taxa. Therefore, taxonomic classification of novel putative MTB species based on short-reads is a challenge (Johnson et al., 2019) and longer read sequencing technology like the circular consensus sequencing (CCS) (Wenger et al., 2019) is required to provide a higher taxonomic accuracy.

Since magnetic enrichment of MTB was not possible, the process of discovering novel MTB taxa is limited by sequence homology to a database and requires further evidence. Taxonomic classification of the recovered MTB genomes using bacterial marker genes used for phylogenetic inference (Chaumeil et al., 2019) showed the genomes to be affiliated with the family *Magnetovibrionaceae* and *Magnetospiraceae*. The most resolved classification was at the genus *Magnetovibrio*, thus supporting the presence of novel putative MTB species based on short read sequencing. SCG analysis also revealed *Magnetovibrio blakemorei* to be the closest genome from the Genome Taxonomy Database (GTDB), having an average nucleotide identity of less than 80%. This is supported by phylogenetic analysis which showed that bins 234_sub, bin.1389 and bin.1085 were present in the same clade as *Magnetovibrio blakemorei*. This provides another line of evidence that the recovered MTB genomes are from novel putative MTB species.

When contigs were annotated with the KEGG and Gene Ontology (GO) database, contigs could not be annotated to genes (e.g. *mam* and *mad* genes) in the MGCs that play important roles in the formation of functional magnetosomes (Schuler,

671 2008; Murat et al., 2010), except for a large abundance of genes encoding for MCPs.
672 Even when contigs were compared with a database of *mam* genes constructed from
673 UniProt database, the alignment length of genes was shorter than the length of the
674 *mam* genes. To further probe for the presence of mam genes in the metagenomes,
675 HMM-comparisons were used to search for homologs to the *mam* genes in the MTB
676 bins. HMM-comparisons have enhanced sensitivity and alignment quality over
677 sequence-based alignments. The three bins most similar to *Magnetovibrio*
678 *blakemorei* did not have exact matches to the *mam* genes but have proteins that are
679 homologous to them. The identified proteins have more than 50% probability of
680 being homologs and could likely play a similar function to the *mam* genes, though
681 further validation is required.

682 Multiple lines of evidence through bioinformatic analyses have shown MTB to be
683 present in our samples. The genomes recovered are phylogenetically related to
684 MTB, but they do not possess the full-length MCG genes (e.g. *mam* and *mad* genes)
685 involved in magnetosome formation nor the magnetotactic behaviour. In some
686 axenic cultures of MTB, the loss of the magnetic phenotype can be easily induced
687 through a deletion of the MGC (Schubbe et al., 2003; Ullrich et al., 2005) or a
688 change in environmental conditions (Urban, 2000). The lack of magnetotactic
689 behaviour could be explained by the shallow inclination angle of the Earth's magnetic
690 field at our sampling site: -13^{o} and a weaker magnetic field which imparts no
691 advantage in guiding MTB along vertical axis to favourable conditions.
692 Consequently, the lack of selective pressure (Albalat and Canestro, 2016) causes
693 MTB to lose their ability to synthesize magnetosomes through evolution and hence,
694 the accompanying magnetic phenotype (Lefevre and Bazylinski, 2013; Lin et al.,
695 2018).

MTB in natural habitats mainly present at the oxic-anoxic interface (OAI). In our study, the OAI would be at the sediment-seawater interface. Sediment and seawater slurry samples were obtained using a Van Veen grab. During the sampling process, the spatial stratification was disrupted, and the slurry samples were homogenized. DNA extraction from all the collected slurry sample (in the order of kg) would practically not be feasible while extraction from 0.5 g of the slurry sample would not be representative for profiling MTB communities. Hence, we incubate the slurry samples for an extended period of time to re-establish the OAI and before we collect the slurry samples at the OAI and sequence them. As samples were stored under prolonged incubation and have undergone ecological succession, the diversity of MTB reported is not representative of the original sample. To obtain a true representation of the MTB community in the marine environment, samples maintaining their spatial stratification at the OAI in the natural environments should be sequenced directly after sample collection. In addition, the MTB biodiversity reported in this work was obtained from a small sample size in terms of geographical and temporal variation, and more sampling would be required to depict a more accurate representation of the MTB biodiversity. Another limitation of this study is the lack of water quality dynamics data at the sampling sites. A correlation of MTB biodiversity to water quality would bring in more insights toward understanding the ecophysiology of MTB in the tropical marine environment of Singapore.

This study fills in the knowledge gap of MTB biodiversity in the tropical marine environment near the geo-equator. Our key findings include (i) Majority of the MTB in the tropical marine environment of Singapore represents novel MTB taxa that cannot be classified at the species level; (ii) The relative abundance of MTB (~0.2-1.69%) in the tropical marine environment of Singapore was found to be substantially lower

than studies for other regions; (iii) Genera *Magnetovibrio* and *Desulfamplus*, but not *Magnetococcus*, were the dominant MTB in the tropical marine environment of Singapore; and (iv) Three MTB genomic bins were recovered and they are unclassified at the species level, with *Magnetovibrio blakemorei* being the closest-associated genome. Our findings will facilitate future research efforts aiming to unravel the ecological roles of MTB in the tropical marine environments as well as to bioprospecting novel MTB that have been adapted to tropical marine environments for biotechnological applications.

Acknowledgements

This research was supported by the National Research Foundation (NRF), Prime Minister's Office, Singapore under its Marine Science Research and Development Programme (MSRDP-P30). We thank Dr. Ooi Seng Keat and the Ecological Monitoring, Informatics and Dynamics Group at the Tropical Marine Science Institute (TMSI), National University of Singapore (NUS) and Dr. Federico Lauro and his team from the Asian School of the Environment, Nanyang Technological University for allowing us to join the sampling trips for the collection of marine sediments and seawater. We are grateful for the technical staff from the SCELSE sequencing laboratory for their assistance. The computational work for this article was done on resources of the National Supercomputing Computer, Singapore (https://www.nscc.sg).

References

Abreu, F., Carolina, A., Araujo, V., Leao, P., Silva, K.T., Carvalho, F.M., et al. (2016). Culture-independent characterization of novel psychrophilic magnetotactic cocci from Antarctic marine sediments. *Environ Microbiol* 18(12)**,** 4426-4441. doi: 10.1111/1462-2920.13388.

746 Abreu, F., Morillo, V., Nascimento, F.F., Werneck, C., Cantao, M.E., Ciapina, L.P., et
747 al. (2014). Deciphering unusual uncultured magnetotactic multicellular
748 prokaryotes through genomics. *ISME J* 8(5)**,** 1055-1068. doi:
749 10.1038/ismej.2013.203.
750 Albalat, R., and Canestro, C. (2016). Evolution by gene loss. *Nat Rev Genet* 17(7)**,**
751 379-391. doi: 10.1038/nrg.2016.39.
752 Alexandre, G., Greer-Phillips, S., and Zhulin, I.B. (2004). Ecological role of energy
753 taxis in microorganisms. *FEMS Microbiol Rev* 28(1)**,** 113-126. doi:
754 10.1016/j.femsre.2003.10.003.
755 Alneberg, J., Bjarnason, B.S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., et al.
756 (2014). Binning metagenomic contigs by coverage and composition. *Nat*
757 *Methods* 11(11)**,** 1144-1146. doi: 10.1038/nmeth.3103.
758 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic
759 local alignment search tool. *J Mol Biol* 215(3)**,** 403-410. doi: 10.1016/S0022-
760 2836(05)80360-2.
761 Amor, M., Ceballos, A., Wan, J., Simon, C.P., Aron, A.T., Chang, C.J., et al. (2020).
762 Magnetotactic Bacteria Accumulate a Large Pool of Iron Distinct from Their
763 Magnetite Crystals. *Appl Environ Microbiol* 86(22). doi: 10.1128/AEM.01278-
764 20.
765 Andersen, K.S., Kirkegaard, R.H., Karst, S.M., and Albertsen, M. (2018). ampvis2:
766 an R package to analyse and visualise 16S rRNA amplicon data. *bioRxiv***,**
767 299537. doi: 10.1101/299537.
768 Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data*
769 [Online]. Available: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
770 [Accessed].
771 Arango-Argoty, G., Singh, G., Heath, L.S., Pruden, A., Xiao, W., and Zhang, L.
772 (2016). MetaStorm: A Public Resource for Customizable Metagenomics
773 Annotation. *PLoS One* 11(9)**,** e0162442. doi: 10.1371/journal.pone.0162442.
774 Araujo, F.A., Barh, D., Silva, A., Guimaraes, L., and Ramos, R.T.J. (2018). GO
775 FEAT: a rapid web-based functional annotation tool for genomic and
776 transcriptomic data. *Sci Rep* 8(1)**,** 1794. doi: 10.1038/s41598-018-20211-9.
777 Asnicar, F., Thomas, A.M., Beghini, F., Mengoni, C., Manara, S., Manghi, P., et al.
778 (2020). Precise phylogenetic analysis of microbial isolates and genomes from
779 metagenomes using PhyloPhlAn 3.0. *Nat Commun* 11(1)**,** 2500. doi:
780 10.1038/s41467-020-16366-7.
781 Baker, A.E., Diepold, A., Kuchma, S.L., Scott, J.E., Ha, D.G., Orazi, G., et al. (2016).
782 PilZ Domain Protein FlgZ Mediates Cyclic Di-GMP-Dependent Swarming
783 Motility Control in Pseudomonas aeruginosa. *J Bacteriol* 198(13)**,** 1837-1846.
784 doi: 10.1128/JB.00196-16.
785 Barber-Zucker, S., Keren-Khadmy, N., and Zarivach, R. (2016). From invagination to
786 navigation: The story of magnetosome-associated proteins in magnetotactic
787 bacteria. *Protein Sci* 25(2)**,** 338-351. doi: 10.1002/pro.2827.
788 Bazylinski, D., Lefèvre, C., and Schüler, D. (2013). "The Prokaryotes."), 453-494.
789 Bazylinski, D.A., and Frankel, R.B. (2004). Magnetosome formation in prokaryotes.
790 *Nat Rev Microbiol* 2(3)**,** 217-230. doi: 10.1038/nrmicro842.
791 Bengtsson-Palme, J., Hartmann, M., Eriksson, K.M., Pal, C., Thorell, K., Larsson,
792 D.G., et al. (2015). METAXA2: improved identification and taxonomic
793 classification of small and large subunit rRNA in metagenomic data. *Mol Ecol*
794 *Resour* 15(6)**,** 1403-1414. doi: 10.1111/1755-0998.12399.

Bengtsson-Palme, J., Richardson, R.T., Meola, M., Wurzbacher, C., Tremblay, E.D., Thorell, K., et al. (2018). Metaxa2 Database Builder: enabling taxonomic identification from metagenomic or metabarcoding data using any genetic marker. *Bioinformatics* 34(23)**,** 4027-4033. doi: 10.1093/bioinformatics/bty482.

Bennet, M., McCarthy, A., Fix, D., Edwards, M.R., Repp, F., Vach, P., et al. (2014). Influence of magnetic fields on magneto-aerotaxis. *PLoS One* 9(7)**,** e101150. doi: 10.1371/journal.pone.0101150.

Blakemore, R.P., Frankel, R.B., and Kalmijn, A.J. (1980). South-seeking magnetotactic bacteria in the Southern Hemisphere. *Nature* 286(5771)**,** 384-385. doi: 10.1038/286384a0.

Bray, J.R., and Curtis, J.T. (1957). An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* 27(4)**,** 325-349. doi: 10.2307/1942268.

Breitwieser, F.P., and Salzberg, S.L. (2016). Pavian: Interactive analysis of metagenomics data for microbiomics and pathogen identification. *bioRxiv***,** 084715. doi: 10.1101/084715.

Chaumeil, P.A., Mussig, A.J., Hugenholtz, P., and Parks, D.H. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*. doi: 10.1093/bioinformatics/btz848.

Chénard, C., Wijaya, W., Vaulot, D., dos Santos, A.L., Martin, P., Kaur, A., et al. (2019). Temporal dynamics of Bacteria, Archaea and protists in equatorial coastal waters. *bioRxiv***,** 658278. doi: 10.1101/658278.

Descamps, E.C.T., Monteil, C.L., Menguy, N., Ginet, N., Pignol, D., Bazylinski, D.A., et al. (2017). Desulfamplus magnetovallimortis gen. nov., sp. nov., a magnetotactic bacterium from a brackish desert spring able to biomineralize greigite and magnetite, that represents a novel lineage in the Desulfobacteraceae. *Syst Appl Microbiol* 40(5)**,** 280-289. doi: 10.1016/j.syapm.2017.05.001.

Dong, Y., Li, J., Zhang, W., Zhang, W., Zhao, Y., Xiao, T., et al. (2016). The detection of magnetotactic bacteria in deep sea sediments from the east Pacific Manganese Nodule Province. *Environ Microbiol Rep* 8(2)**,** 239-249. doi: 10.1111/1758-2229.12374.

Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., et al. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3**,** e1319. doi: 10.7717/peerj.1319.

Faivre, D., and Schuler, D. (2008). Magnetotactic bacteria and magnetosomes. *Chem Rev* 108(11)**,** 4875-4898. doi: 10.1021/cr078258w.

Falke, J.J., and Hazelbauer, G.L. (2001). Transmembrane signaling in bacterial chemoreceptors. *Trends Biochem Sci* 26(4)**,** 257-265.

Flies, C.B., Jonkers, H.M., de Beer, D., Bosselmann, K., Bottcher, M.E., and Schuler, D. (2005a). Diversity and vertical distribution of magnetotactic bacteria along chemical gradients in freshwater microcosms. *FEMS Microbiol Ecol* 52(2)**,** 185-195. doi: 10.1016/j.femsec.2004.11.006.

Flies, C.B., Peplies, J., and Schuler, D. (2005b). Combined approach for characterization of uncultivated magnetotactic bacteria from various aquatic environments. *Appl Environ Microbiol* 71(5)**,** 2723-2731. doi: 10.1128/AEM.71.5.2723-2731.2005.

Frankel, R.B., Bazylinski, D.A., Johnson, M.S., and Taylor, B.L. (1997). Magneto-aerotaxis in marine coccoid bacteria. *Biophys J* 73(2)**,** 994-1000. doi: 10.1016/S0006-3495(97)78132-3.

Frankel, R.B., Blakemore, R.P., De Araujo, F.F.T., Esquivel, D.M.S., and Danon, J. (1981). Magnetotactic Bacteria at the Geomagnetic Equator. *Science* 212(4500)**,** 1269. doi: 10.1126/science.212.4500.1269.

Gans, J., Wolinsky, M., and Dunbar, J. (2005). Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science* 309(5739)**,** 1387-1390. doi: 10.1126/science.1112665.

Hazelbauer, G.L., Falke, J.J., and Parkinson, J.S. (2008). Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem Sci* 33(1)**,** 9-19. doi: 10.1016/j.tibs.2007.09.014.

Heyen, U., and Schuler, D. (2003). Growth and magnetosome formation by microaerophilic Magnetospirillum strains in an oxygen-controlled fermentor. *Appl Microbiol Biotechnol* 61(5-6)**,** 536-544. doi: 10.1007/s00253-002-1219-x.

Huson, D.H., Beier, S., Flade, I., Gorska, A., El-Hadidi, M., Mitra, S., et al. (2016). MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Comput Biol* 12(6)**,** e1004957. doi: 10.1371/journal.pcbi.1004957.

Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11**,** 119. doi: 10.1186/1471-2105-11-119.

Jansson, J.K., Neufeld, J.D., Moran, M.A., and Gilbert, J.A. (2012). Omics for understanding microbial functional dynamics. *Environ Microbiol* 14(1)**,** 1-3. doi: 10.1111/j.1462-2920.2011.02518.x.

Jogler, C., and Schuler, D. (2009). Genomics, genetics, and cell biology of magnetosome formation. *Annu Rev Microbiol* 63**,** 501-521. doi: 10.1146/annurev.micro.62.081307.162908.

Johnson, J.S., Spakowicz, D.J., Hong, B.-Y., Petersen, L.M., Demkowicz, P., Chen, L., et al. (2019). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications* 10(1)**,** 5029. doi: 10.1038/s41467-019-13036-1.

Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *J Mol Biol* 428(4)**,** 726-731. doi: 10.1016/j.jmb.2015.11.006.

Kang, D.D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7**,** e7359. doi: 10.7717/peerj.7359.

Keim, C.N., Solorzano, G., Farina, M., and Lins, U. (2005). Intracellular inclusions of uncultured magnetotactic bacteria. *Int Microbiol* 8(2)**,** 111-117.

Kim, D., Song, L., Breitwieser, F.P., and Salzberg, S.L. (2016). Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Res* 26(12)**,** 1721-1729. doi: 10.1101/gr.210641.116.

Kolinko, S., Jogler, C., Katzmann, E., Wanner, G., Peplies, J., and Schüler, D. (2012). Single-cell analysis reveals a novel uncultivated magnetotactic bacterium within the candidate division OP3. *Environmental Microbiology* 14(7)**,** 1709-1721. doi: 10.1111/j.1462-2920.2011.02609.x.

Kolinko, S., Richter, M., Glockner, F.O., Brachmann, A., and Schuler, D. (2016). Single-cell genomics of uncultivated deep-branching magnetotactic bacteria reveals a conserved set of magnetosome genes. *Environ Microbiol* 18(1)**,** 21-37. doi: 10.1111/1462-2920.12907.

Kolinko, S., Wanner, G., Katzmann, E., Kiemer, F., Fuchs, B.M., and Schuler, D. (2013). Clone libraries and single cell genome amplification reveal extended

diversity of uncultivated magnetotactic bacteria from marine and freshwater environments. *Environ Microbiol* 15(5)**,** 1290-1301. doi: 10.1111/1462-2920.12004.

Lefevre, C.T., and Bazylinski, D.A. (2013). Ecology, diversity, and evolution of magnetotactic bacteria. *Microbiol Mol Biol Rev* 77(3)**,** 497-526. doi: 10.1128/MMBR.00021-13.

Lefevre, C.T., Frankel, R.B., Abreu, F., Lins, U., and Bazylinski, D.A. (2011a). Culture-independent characterization of a novel, uncultivated magnetotactic member of the Nitrospirae phylum. *Environ Microbiol* 13(2)**,** 538-549. doi: 10.1111/j.1462-2920.2010.02361.x.

Lefevre, C.T., Menguy, N., Abreu, F., Lins, U., Posfai, M., Prozorov, T., et al. (2011b). A cultured greigite-producing magnetotactic bacterium in a novel group of sulfate-reducing bacteria. *Science* 334(6063)**,** 1720-1723. doi: 10.1126/science.1212596.

Lefevre, C.T., Trubitsyn, D., Abreu, F., Kolinko, S., Jogler, C., de Almeida, L.G., et al. (2013). Comparative genomic analysis of magnetotactic bacteria from the Deltaproteobacteria provides new insights into magnetite and greigite magnetosome genes required for magnetotaxis. *Environ Microbiol* 15(10)**,** 2712-2735. doi: 10.1111/1462-2920.12128.

Lefevre, C.T., Viloria, N., Schmidt, M.L., Posfai, M., Frankel, R.B., and Bazylinski, D.A. (2012). Novel magnetite-producing magnetotactic bacteria belonging to the Gammaproteobacteria. *ISME J* 6(2)**,** 440-450. doi: 10.1038/ismej.2011.97.

Lefevre, C.T., and Wu, L.F. (2013). Evolution of the bacterial organelle responsible for magnetotaxis. *Trends Microbiol* 21(10)**,** 534-543. doi: 10.1016/j.tim.2013.07.005.

Li, D., Luo, R., Liu, C.M., Leung, C.M., Ting, H.F., Sadakane, K., et al. (2016). MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* 102**,** 3-11. doi: 10.1016/j.ymeth.2016.02.020.

Lin, W., Bazylinski, D.A., Xiao, T., Wu, L.F., and Pan, Y. (2014). Life with compass: diversity and biogeography of magnetotactic bacteria. *Environ Microbiol* 16(9)**,** 2646-2658. doi: 10.1111/1462-2920.12313.

Lin, W., Li, J., and Pan, Y. (2012a). Newly isolated but uncultivated magnetotactic bacterium of the phylum Nitrospirae from Beijing, China. *Appl Environ Microbiol* 78(3)**,** 668-675. doi: 10.1128/AEM.06764-11.

Lin, W., Li, J., Schuler, D., Jogler, C., and Pan, Y. (2009). Diversity analysis of magnetotactic bacteria in Lake Miyun, northern China, by restriction fragment length polymorphism. *Syst Appl Microbiol* 32(5)**,** 342-350. doi: 10.1016/j.syapm.2008.10.005.

Lin, W., and Pan, Y. (2015). A putative greigite-type magnetosome gene cluster from the candidate phylum Latescibacteria. *Environ Microbiol Rep* 7(2)**,** 237-242. doi: 10.1111/1758-2229.12234.

Lin, W., Pan, Y., and Bazylinski, D.A. (2017a). Diversity and ecology of and biomineralization by magnetotactic bacteria. *Environ Microbiol Rep* 9(4)**,** 345-356. doi: 10.1111/1758-2229.12550.

Lin, W., Paterson, G.A., Zhu, Q., Wang, Y., Kopylova, E., Li, Y., et al. (2017b). Origin of microbial biomineralization and magnetotaxis during the Archean. *Proc Natl Acad Sci U S A* 114(9)**,** 2171-2176. doi: 10.1073/pnas.1614654114.

943 Lin, W., Wang, Y., Gorby, Y., Nealson, K., and Pan, Y. (2013). Integrating niche-
944 based process and spatial process in biogeography of magnetotactic bacteria.
945 *Sci Rep* 3**,** 1643. doi: 10.1038/srep01643.
946 Lin, W., Wang, Y., Li, B., and Pan, Y. (2012b). A biogeographic distribution of
947 magnetotactic bacteria influenced by salinity. *ISME J* 6(2)**,** 475-479. doi:
948 10.1038/ismej.2011.112.
949 Lin, W., Zhang, W., Zhao, X., Roberts, A.P., Paterson, G.A., Bazylinski, D.A., et al.
950 (2018). Genomic expansion of magnetotactic bacteria reveals an early
951 common origin of magnetotaxis with lineage-specific evolution. *ISME J* 12(6)**,**
952 1508-1519. doi: 10.1038/s41396-018-0098-9.
953 Lins, U., and Bazylinski, D.A. (2009). "Magnetotaxis," in *Encyclopedia of*
954 *Microbiology (Third Edition),* ed. M. Schaechter. (Oxford: Academic Press),
955 229-241.
956 Liu, J., Zhang, W., Li, X., Li, X., Chen, X., Li, J.H., et al. (2017). Bacterial community
957 structure and novel species of magnetotactic bacteria in sediments from a
958 seamount in the Mariana volcanic arc. *Sci Rep* 7(1)**,** 17964. doi:
959 10.1038/s41598-017-17445-4.
960 Martins, J.L., Silveira, T.S., Abreu, F., de Almeida, F.P., Rosado, A.S., and Lins, U.
961 (2012). Spatiotemporal distribution of the magnetotactic multicellular
962 prokaryote Candidatus Magnetoglobus multicellularis in a Brazilian
963 hypersaline lagoon and in microcosms. *Int Microbiol* 15(3)**,** 141-149. doi:
964 10.2436/20.1501.01.167.
965 Matsunaga, T., Okamura, Y., Fukuda, Y., Wahyudi, A.T., Murase, Y., and
966 Takeyama, H. (2005). Complete genome sequence of the facultative
967 anaerobic magnetotactic bacterium Magnetospirillum sp. strain AMB-1. *DNA*
968 *Res* 12(3)**,** 157-166. doi: 10.1093/dnares/dsi002.
969 Murat, D., Quinlan, A., Vali, H., and Komeili, A. (2010). Comprehensive genetic
970 dissection of the magnetosome gene island reveals the step-wise assembly of
971 a prokaryotic organelle. *Proc Natl Acad Sci U S A* 107(12)**,** 5593-5598. doi:
972 10.1073/pnas.0914439107.
973 NOAA (2019a). *US/UK World Magnetic Model - 2019, Main Field Down Component*
974 *(Z)* [Online]. Available:
975 https://www.ngdc.noaa.gov/geomag/WMM/data/WMM2015/WMM2015v2_Z_
976 MERC.pdf [Accessed].
977 NOAA (2019b). *US/UK World Magnetic Model - 2019, Main Field Inclination (I)*
978 [Online]. Available:
979 https://www.ngdc.noaa.gov/geomag/WMM/data/WMM2015/WMM2015v2_I_M
980 ERC.pdf [Accessed].
981 Oestreicher, Z., Lower, S.K., Lin, W., and Lower, B.H. (2012). Collection, isolation
982 and enrichment of naturally occurring magnetotactic bacteria from the
983 environment. *J Vis Exp* (69)**,** e50123. doi: 10.3791/50123.
984 Oksansen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et
985 al. (2019). Vegan: Community ecology package. R package version 2.5-5.
986 Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015).
987 CheckM: assessing the quality of microbial genomes recovered from isolates,
988 single cells, and metagenomes. *Genome Research* 25(7)**,** 1043-1055. doi:
989 10.1101/gr.186072.114.
990 Postec, A., Tapia, N., Bernadac, A., Joseph, M., Davidson, S., Wu, L.F., et al.
991 (2012). Magnetotactic bacteria in microcosms originating from the French

Mediterranean Coast subjected to oil industry activities. *Microb Ecol* 63(1)**,** 1-11. doi: 10.1007/s00248-011-9910-z.

Pratt, J.T., Tamayo, R., Tischler, A.D., and Camilli, A. (2007). PilZ domain proteins bind cyclic diguanylate and regulate diverse processes in Vibrio cholerae. *J Biol Chem* 282(17)**,** 12860-12870. doi: 10.1074/jbc.M611593200.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41(Database issue)**,** D590-596. doi: 10.1093/nar/gks1219.

Ryjenkov, D.A., Simm R Fau - Romling, U., Romling U Fau - Gomelsky, M., and Gomelsky, M. (2006). The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ domain protein YcgR controls motility in enterobacteria. (0021-9258 (Print)).

Sakaguchi, T., Tsujimura, N., and Matsunaga, T. (1996). A novel method for isolation of magnetic bacteria without magnetic collection using magnetotaxis. *Journal of Microbiological Methods* 26(1-2)**,** 139-145. doi: Doi 10.1016/0167-7012(96)00905-0.

Schubbe, S., Kube, M., Scheffel, A., Wawer, C., Heyen, U., Meyerdierks, A., et al. (2003). Characterization of a spontaneous nonmagnetic mutant of Magnetospirillum gryphiswaldense reveals a large deletion comprising a putative magnetosome island. *J Bacteriol* 185(19)**,** 5779-5790. doi: 10.1128/jb.185.19.5779-5790.2003.

Schuler, D. (1999). Formation of magnetosomes in magnetotactic bacteria. *J Mol Microbiol Biotechnol* 1(1)**,** 79-86.

Schuler, D. (2002). The biomineralization of magnetosomes in Magnetospirillum gryphiswaldense. *Int Microbiol* 5(4)**,** 209-214. doi: 10.1007/s10123-002-0086-8.

Schuler, D. (2008). Genetics and cell biology of magnetosome formation in magnetotactic bacteria. *FEMS Microbiol Rev* 32(4)**,** 654-672. doi: 10.1111/j.1574-6976.2008.00116.x.

Sieber, C.M.K., Probst, A.J., Sharrar, A., Thomas, B.C., Hess, M., Tringe, S.G., et al. (2018). Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol* 3(7)**,** 836-843. doi: 10.1038/s41564-018-0171-1.

Simmons, S.L., Bazylinski, D.A., and Edwards, K.J. (2007). Population dynamics of marine magnetotactic bacteria in a meromictic salt pond described with qPCR. *Environ Microbiol* 9(9)**,** 2162-2174. doi: 10.1111/j.1462-2920.2007.01330.x.

Simmons, S.L., Sievert, S.M., Frankel, R.B., Bazylinski, D.A., and Edwards, K.J. (2004). Spatiotemporal distribution of marine magnetotactic bacteria in a seasonally stratified coastal salt pond. *Appl Environ Microbiol* 70(10)**,** 6230-6239. doi: 10.1128/AEM.70.10.6230-6239.2004.

Spring, S., Amann, R., Ludwig, W., Schleifer, K.-H., Schüler, D., Poralla, K., et al. (1995). Phylogenetic Analysis of Uncultured Magnetotactic Bacteria from the Alpha-Subclass of Proteobacteria. *Systematic and Applied Microbiology* 17(4)**,** 501-508. doi: https://doi.org/10.1016/S0723-2020(11)80068-8.

Spring, S., Amann, R., Ludwig, W., Schleifer, K.H., van Gemerden, H., and Petersen, N. (1993). Dominating role of an unusual magnetotactic bacterium in the microaerobic zone of a freshwater sediment. *Appl Environ Microbiol* 59(8)**,** 2397-2403.

Steinegger, M., Meier, M., Mirdita, M., Vohringer, H., Haunsberger, S.J., and Soding, J. (2019). HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics* 20(1)**,** 473. doi: 10.1186/s12859-019-3019-7.

Tan, K.S., Acerbi, E., and Lauro, F.M. (2016). Marine habitats and biodiversity of Singapore's coastal waters: A review. *Regional Studies in Marine Science* 8**,** 340-352. doi: https://doi.org/10.1016/j.rsma.2016.01.008.

Tully, B.J., Graham, E.D., and Heidelberg, J.F. (2018). The reconstruction of 2,631 draft metagenome-assembled genomes from the global oceans. *Sci Data* 5**,** 170203. doi: 10.1038/sdata.2017.203.

Ullrich, S., Kube, M., Schubbe, S., Reinhardt, R., and Schuler, D. (2005). A hypervariable 130-kilobase genomic region of Magnetospirillum gryphiswaldense comprises a magnetosome island which undergoes frequent rearrangements during stationary growth. *J Bacteriol* 187(21)**,** 7176-7184. doi: 10.1128/JB.187.21.7176-7184.2005.

UniProt, C. (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47(D1)**,** D506-D515. doi: 10.1093/nar/gky1049.

Urban, J.E. (2000). Adverse effects of microgravity on the magnetotactic bacterium Magnetospirillum magnetotacticum. *Acta Astronaut* 47(10)**,** 775-780. doi: 10.1016/s0094-5765(00)00120-x.

Wenger, A.M., Peluso, P., Rowell, W.J., Chang, P.C., Hall, R.J., Concepcion, G.T., et al. (2019). Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nat Biotechnol* 37(10)**,** 1155-1162. doi: 10.1038/s41587-019-0217-9.

Wenter, R., Wanner, G., Schüler, D., and Overmann, J. (2009). Ultrastructure, tactic behaviour and potential for sulfate reduction of a novel multicellular magnetotactic prokaryote from North Sea sediments. *Environmental Microbiology* 11(6)**,** 1493-1505. doi: 10.1111/j.1462-2920.2009.01877.x.

Xu, C. (2018). Distribution and diversity of magnetotactic bacteria in sediments of the Yellow Sea continental shelf. *Journal of soils and sediments* v. 18(no. 7)**,** pp. 13-2646-2018 v.2618 no.2647. doi: 10.1007/s11368-018-1912-8.

Zhang, R., Chen, Y.R., Du, H.J., Zhang, W.Y., Pan, H.M., Xiao, T., et al. (2014). Characterization and phylogenetic identification of a species of spherical multicellular magnetotactic prokaryotes that produces both magnetite and greigite crystals. *Res Microbiol* 165(7)**,** 481-489. doi: 10.1016/j.resmic.2014.07.012.